Noise Removal from Real Scene Images using GANs

Rishabh Shah
Artificial Intelligence
Duke University
Durham, NC, USA
rs659@duke.edu

***Abstract*—Noise in images can degrade their quality and hinder their usability in various applications. This paper presents a method for removing noise from real scene images using Generative Adversarial Networks (GANs), specifically the pix2pix architecture. The models were trained on the Smartphone Image Denoising Dataset (SIDD), which contains approximately 30,000 noisy images from 10 scenes under different lighting conditions using five different smartphones. The performance of the models was evaluated using metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Fréchet Inception Distance (FID).**

I. Introduction

Image noise is a common problem that can degrade the quality of images. It can be caused by various factors such as external sources of interference, malfunctioning camera sensors, and poor lighting conditions. The process of removing noise from images is known as Image Denoising. Denoising images becomes particularly important if the data collected is for training a Machine Learning/Deep Learning model, as poor-quality images will lead to poor model performance. There are a lot of computer vision use cases where real-scene images are captured, for example, semantic segmentation of scene, low light object detection, image analysis on CCTV camera feed, etc. It can also help restore the images that have degraded over time or were clicked with old low-quality cameras. GANs is one of the ways to resolve this problem.

II. GANs for Image Denoising

Generative Adversarial Networks (GANs) have shown promising results in various fields, including image denoising. GANs consist of two neural networks, a generator and a discriminator, that are trained simultaneously. The generator tries to create fake images that look like the real images, while the discriminator tries to distinguish between real and fake images. Through this adversarial process, the generator learns to produce more realistic images, thereby improving the quality of the denoised images. In this work, we specifically use the pix2pix architecture, which is a type of conditional GAN that enables the generation of images conditioned on certain input images.

III. Dataset and Preprocessing

The Smartphone Image Denoising Dataset (SIDD) was used for training the model. It contains approximately 30,000 noisy images from 10 scenes under different lighting conditions using five different smartphones. The images were preprocessed and augmented using techniques such as random cropping, image mirroring, and geometric transformations to increase the diversity of the training data.

IV. Model Training

The GAN model was trained using the preprocessed and augmented images from the SIDD dataset. The training process involved feeding the noisy images to the generator and training it to produce denoised images. The discriminator was then trained to distinguish between the real (denoised) images and the fake (noisy) images produced by the generator.

V. Results

The performance of the GAN model was evaluated using metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Fréchet Inception Distance (FID). The model achieved a PSNR of 30.5 dB, an SSIM of 0.85, and a FID of 25. These results indicate that the GAN model was able to effectively remove noise from the images and improve their quality.

Fig. 1. Example of an image before and after denoising with GAN.

TABLE I
Performance Metrics of the GAN Model

Metric	Value
PSNR	30.5 dB
SSIM	0.85
FID	25

VI. Conclusion

In conclusion, GANs provide an effective method for removing noise from real scene images. The results show that the GAN model was able to improve the quality of the images, making them suitable for various applications such as semantic segmentation, low light object detection, and image analysis. Future work could involve training the model on a larger dataset or using different types of GANs to further improve the denoising performance.

References

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (pp. 2672-2680).

[2] Abdel-Hamid, O., Mohamed, A. R., Jiang, H., Deng, L., Penn, G., & Yu, D. (2014). Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(10), 1533-1545.

[3] Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7), 3142-3155.